



























# The Data Analysis System Help

-  [System Overview](#)
-  [System Requirements](#)
-  [How to Start DASW](#)
-  [How to Change the System Setup](#)
-  [System Limitations](#)
-  [The DAS Main Window](#)
-  [How to View Description/Percentage](#)
-  [How to Navigate Through Variables](#)
-  [How to Search for Text](#)
-  [How to Create a List of Parameters](#)
-  [How to Resize Parameter List](#)
-  [How to Edit a Parameter](#)
-  [How to Delete a Parameter](#)
-  [How to Create a New Parameter File](#)
-  [How to Import a Parameter File](#)
-  [How to Edit a Parameter File](#)
-  [How to Edit Titles](#)
-  [How to Save a Parameter File](#)
-  [How to Insert a Row](#)
-  [How to Delete a Row](#)
-  [How to Get Tables](#)
-  [How to View DAS Output](#)
-  [How to Print Variable Information](#)
-  [How Do I ?](#)
-  [All About Tags](#)
-  [Percent>](#)

- 👉 [Mean](#)
- 👉 [Average>o](#)
- 👉 [Centile](#)
- 👉 [Centile>0](#)
- 👉 [Column\\_Cat](#)
- 👉 [Column\\_Lump](#)
- 👉 [Column\\_Cut](#)
- 👉 [Row\\_Cat](#)
- 👉 [Row\\_Lump](#)
- 👉 [Row\\_Cut](#)
- 👉 [By\\_Cat](#)
- 👉 [By\\_Lump](#)
- 👉 [By\\_Cut](#)
- 👉 [Spanner\\_Cat](#)
- 👉 [Spanner\\_Lump](#)
- 👉 [Spanner\\_Cut](#)
- 👉 [Filter](#)
- 👉 [Or\\_Filter](#)
- 👉 [Weight](#)
- 👉 [Continuous](#)
- 👉 [Cut](#)
- 👉 [Each](#)
- 👉 [Lump](#)

# System Overview

The Data Analysis System for Windows (DASW) is a software tool which allows the user to browse through the lists of variables and to select variables to produce tables and correlation matrices. There is a separate DASW for each dataset, and all have a consistent interface and command structure.

The program opens with a Main window that lists all of the variables in the selected dataset. Using either a mouse or menu command, you then “tag” or select those variables you want in the table or correlation matrix.

The process of tagging lets you specify in detail how the variables are to be defined and how they will be used to create your table or correlation matrix. For example, you may specify that one variable will appear with each of its values as a separate row in your table and that another variable will appear as a column of its mean values. When all variables have been selected, defined, and placed as rows or columns, select the SAVE option from the File menu. The DASW then produces a text file that contains the information, or parameters, used to compute the request.

## General information about VARIABLES

There are **Categorical** and **Continuous** variables in every DASW system. For both types of variables, values of -1 indicate useless data due to missing values, refusals, legitimate skips, and other factors. Categorical variables generally have nonadditive values—for example, GENDER {Male, Female}, RACE {Black, White, Hispanic, Asian, American Indian}, and MARITAL {Never married, Married, Divorced/separated, Widowed} all have values that should not be added together. A few categorical variables may have special or restricted code values of zero. For example, for the variable PSE1ST {Type of first postsecondary institution}, the value zero indicates that the sample member never entered any postsecondary institution. Continuous variables are additive. Examples include AGE, INCOME, and SES. /\* The above variables may not be in the DAS you are using. \*/

## General Information about TABLES

Tables consist of six key elements: Title, Subgroup specifications, Columns, Column Groupings, Rows, and Row Groupings. The DASW system allows you to specify all of these elements.

Table 1.--Average socioeconomic status among 1992 high school graduates who were not married by 1994, by postsecondary enrollment and gender

And Filters			
Marital status in 1994	=	Never married	
=====			
		---Gender---	
		Male	Female
		Total	
-----			
	----- First PSE type = 4-year -----		
Total		67.7	66.6
			67.9
Race/ethnicity			
Asian or Pacific Islander		66.6	65.3
			66.6
Hispanic regardless of race		47.3	42.9
			47.6
Black not of Hispanic origin		53.3	52.4
			53.4
White not of Hispanic origin		71.4	70.5
			71.5
American Indian or Alaskan native		56.2	low n
			57.0
----- First PSE type = Public 2 year -----			
Total		52.8	50.1
			52.7
Race/ethnicity			

Asian or Pacific Islander	52.1	45.6	52.1
Hispanic regardless of race	36.4	38.3	36.3
Black not of Hispanic origin	38.9	33.7	38.6
White not of Hispanic origin	57.5	55.1	57.6
American Indian or Alaskan native	low n	low n	low n

----- First PSE type = Other -----  
Total 47.4 43.8 47.5

#### Race/ethnicity

Asian or Pacific Islander	low n	low n	low n
Hispanic regardless of race	22.8	21.5	22.6
Black not of Hispanic origin	38.9	31.3	38.9
White not of Hispanic origin	54.3	52.1	54.4
American Indian or Alaskan native	low n	low n	low n

----- First PSE type = Never attended -----  
Total 39.7 38.3 39.7

#### Race/ethnicity

Asian or Pacific Islander	46.6	52.8	49.8
Hispanic regardless of race	27.3	22.6	25.9
Black not of Hispanic origin	31.6	29.8	31.7
White not of Hispanic origin	43.1	42.7	43.3
American Indian or Alaskan native	30.0	26.4	29.9

/

Source: NCES, National Education Longitudinal Study: 1988/94

Computation by DAS-T Version 3.0 on 12/05/1997

**Title.** Government Printing Office (GPO) Standards for table titles require titles to begin with:

*Table k.--Text*

where k is the number of the table within the report. The text may NOT extend beyond the table's column headings. Subsequent lines must be indented to match the beginning of the text on the initial title line. "Among" in table titles identify the subgroup specifications; "by" in table titles identify separate sets of estimates within the table; and "for" in table titles identify general study definitions (e.g., for 1992 high school graduates). Title lines are input by you when you save the parameter file.

**Subgroup specifications.** When a table's estimates are not calculated using all of the cases, there must be a specification of the subgroup used. The description of the subgroup is usually included in the title, but multiple criteria may be included as a footnote. In the DAS, subgroup specifications are defined based on FILTERS (and in some cases WEIGHTS). You may define several FILTERS, but you may not define more than one OR\_FILTER. In table 1, marital status is a FILTER variable.

**Columns.** Columns are sets of estimates where the denominator is defined by the row specification. Columns may be averages, percentages, or centile values. Averages may be defined in the DAS with Mean {including ZEROs as valid data} or Average>0 {including only positive values as valid data} specifications for continuous variables. In table 1, SES is the column variable, with average SES centiles calculated as the estimates.

**Percentages.** The two types of percentages specified within the DASW are single percentage columns and sets of columns of percentages that sum to 100. Single Percent> columns are specified with a cut-point; where the cut-point defines the assignment to the numerator. For example, a Percent> specification for INCOME with a cut-point of 74999.5 will produce estimates of the percentage of cases with incomes of

\$75,000 or more. One of the most used cut-points for Percent> specifications is 0.5, which yields the percentage that had non-zero values for student aid variables, or the percentage receiving aid. Columns of percentages that sum to 100 are specified with Column\_Cat, Column\_Lump, and Column\_Cut specifications.

Centile values for the 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup> (median), 75<sup>th</sup>, and 90<sup>th</sup> percentiles can be generated with the Centile and Centile>0 specification.

**Column Groupings.** In the DASW, single columns of Mean, Average>0, and Percent> specifications may be grouped with By\_Cat, By\_Lump, and By\_Cut specifications. Column\_Cat, Column\_Lump, Column\_Cut, Centile, and Centile>0 may NOT be used with BY variables. BY specifications create subsets of estimates for each category of the BY variable. In table 1, gender is the BY variable.

**Rows.** Rows, including the row variable labels and the row variable categories, form the “stubs” of the rows in the table and define the denominator of the estimate in the table. Rows are specified in the DASW using Row\_Cat, Row\_Lump, and Row\_Cut specifications. In table 1, race/ethnicity is the Row\_Cat variable.

**Row Groupings.** In the DASW, sets of rows are grouped with Spanner\_Cat, Spanner\_Lump, and Spanner\_Cut specifications. In table 1, type of postsecondary education (PSE) institution first attended is the spanner variable.

### **General Information about CORRELATIONS**

Correlation matrices for SPSS or SAS may be created by the DAS. The output file from the correlation version of the DAS is always RMATRIX.PRN. It contains SPSS or SAS code for matrix input of the correlations and design effect information needed to adjust t-tests.

The SETUP.EXE program does NOT install the correlation procedures as an icon in the Data Analysis System group. To set up correlations, see How to Start DASW.

Table and correlation matrices share Titles, Filters, Or\_Filters, and Weights. These parameters label output and define the size of the samples used to calculate the pairwise correlations. The remaining parameters define variables in the matrix.

All correlations in the matrix are calculated **pairwise**. Hence the definition of missing values is more flexible for correlations than for tables. All DAS data have -1 as “missing”. However, some variables have reserved codes of zero. In Cut specifications the set of missing values may include -1 AND 0.

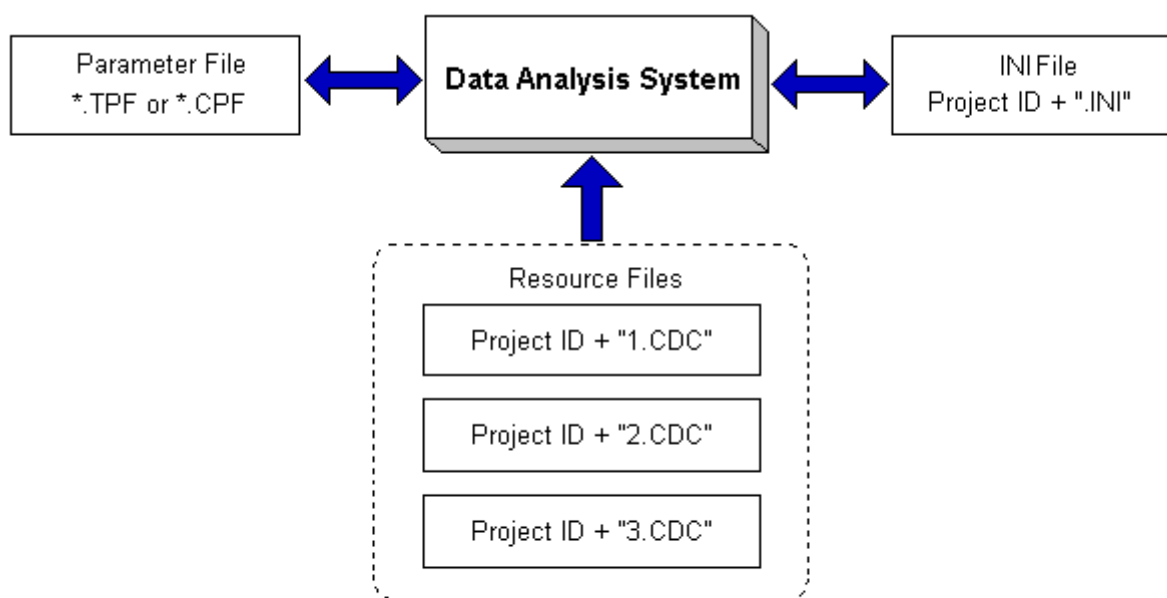
Continuous variables in the matrix create one row and column in the matrix. A variable like AGE with no reserved codes will usually be added to the matrix with “missing” defined as the standard -1. For a months married variable, where zero indicates “Never married”, the missing specification of -1 is not enough. Users need to Filter out the zero values.

By using a Cut tag, a continuous variable may be placed into the correlation matrix as a set of dummy variables {-1,0,1}. After defining “missing”, a range of values is specified for the value ONE--all values outside this range are assigned a value of zero.

Categorical variables may be included with two tags. The Each tag creates a set of dummy variables for each of the non-missing categories except one. Sets of categories may be used to define ONE in a dummy variable by using a Lump tag.

Due to length of string limits in SPSS and SAS, users will probably need to edit all the labels used in the correlation matrix.

# System Requirements



For the successful operation of the Data Analysis System, you must have DASW.EXE, three resource files and an INI file. The project ID determines the names of the resource files and the INI file. The project ID is an abbreviation of the full project name.

For example, if the full project name is "87 Graduate and First-Professional Students", the project ID is "N7G". Resource file names for this project are "N7G1.CDC", "N7G2.CDC", "N7G3.CDC". The INI file name is "N7G.INI".

A typical arrangement of files will have DASW.EXE and N7G.INI in C:\DASW. The N7G?.CDC files will be in C:\DASW\N7G.

See also: How to Start DASW.

# How to Start DASW

After running SETUP.EXE to install the DAS, there should be a folder or group labeled Data Analysis System.

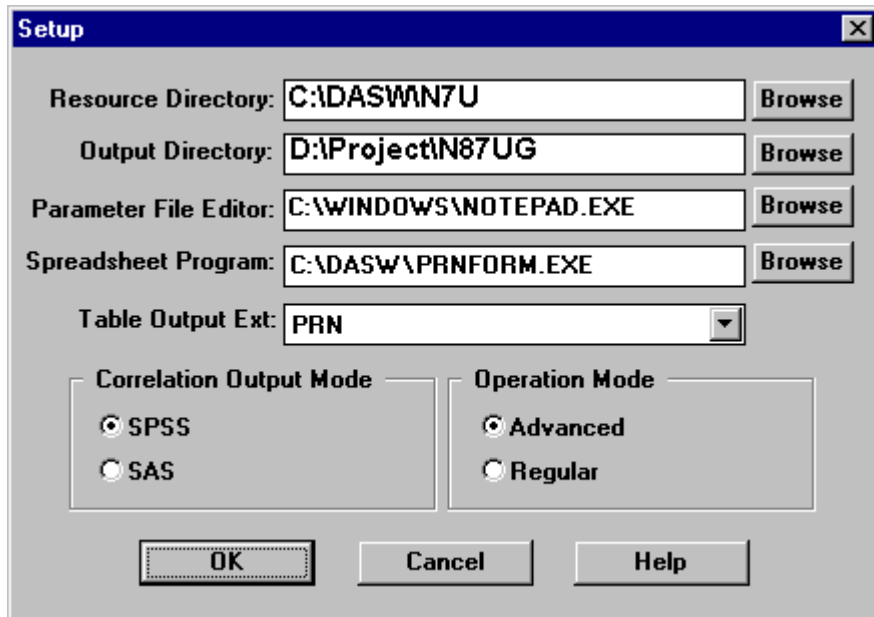
SETUP.EXE only installs the DAS for tables.

To run a table, double click on the icon in the DAS group. You may also use the Windows Run command to start a table. For example, click "**Start**" and then "**Run**". A dialogue box will open where you type in "C:\DASW\DASW T=F94" to run the BPS DAS if you installed the DAS on your C drive in the DASW subdirectory. Click "**OK**" to run your program.

For correlations, you must use the RUN command. For example, to build a NELS correlation matrix, the RUN command would be: **C:\DASW\DASW C=NEL**

# How to Change the System Setup

Choose the **Setup** option from the **File** menu.



## Dialog Box Options

### **Resource Directory**

The name of the directory where the resource files (\*.cdc files) are located. The DASW requires these files in order to operate properly. See System Requirements for details. To change the directory, click on the Browse button to the right. The Windows standard directory dialog box that pops up allows you to choose the directory where the CDC files are located. Click on the OK button to return to the Setup window. Since the resource directory field is read only, you can change the directory only by using the browse button. **Users rarely need to change the resource directory.**

### **Output Directory**

The name of the directory for storage of all output files (e.g., SAS and SPSS code files, FOR CORRELATIONS, tpf AND cpf FILES). To change the directory, click on the Browse button to the right. The Windows standard directory dialog box that pops up allows you to choose the directory you want. Click on the OK button to return to the Setup window. Since the directory field is read only, you can change the directory only by using the browse button. **If you are working on multiple projects you may wish to place .tpf and .cpf files in project-specific directories.**

### **Parameter File Editor**

The name of the file editor with which the parameter file is browsed and updated. To change the file editor, click on the **Browse** button to the right. The Windows standard directory dialog box that pops up allows you to choose the file editor you want. Click on the OK button to return to the Setup window. Since the directory field is read only, you can change the directory only by using the Browse button. The default is NOTEPAD.EXE.

### **Spreadsheet Program**

The name of the table output file viewer with which the table output file is browsed. To change this, click on the **Browse** button to the right. The Windows standard directory dialog box that pops up allows you to choose



the file editor you want. Click on the OK button to return to the Setup window. Since the directory field is read only, you can change the directory only by using the **Browse** button. The default is PRNFORM.EXE. PRNFORM is a routine used to format printed tables. PRNFORM allows users to specify the width of columns, number of decimals displayed for estimates, and formats. PRNFORM works with NOTEPAD.EXE.

### ***Table Output Ext***

The field is available only for the DAS tables. You may select the name of the file extension generated by the DAS tables. To change it, click on the arrow button to the right. Then choose the name you want from the list. Other than the obvious difference in file names, the only difference between .csv and .prn files is the column heading strings. In .csv files, the column headings are written as one long string. In .prn files, the column headings are broken into several rows of strings (the break point character is a blank). Use .csv for Excel, and .prn for Lotus or Quattro.

### ***Correlation Output Mode***

The output mode specifies whether SAS or SPSS correlations are produced.

### ***Operation Mode***

The operation mode which is used to control the flow of the tag operation. When you choose "Advanced," a tag window for Percent>, Mean, Average>0, Weight, Row, Column, and Spanner Category will not be displayed when you create a parameter. Otherwise a tag window displays when you create a parameter, allowing you to change the label if you wish.

### ***OK***

Saves the new directories into the projectid.ini file and returns to the main window.

### ***Cancel***

Cancels the data entered and returns to the main window.

# System Limitations

The system automatically keeps track of its internal size limitations and disables tags or issues error messages. While we believe that the internal limits will probably never be reached by the average user, the following list of all size limits is included for the benefit of advanced users.

There is a fixed limit of 499 tags. The maximum table size is 9,999 cells. The maximum number of columns is 99, and the maximum number of rows (including all row labels) is 999.

There is a maximum of 99 filters, with an additional overall maximum of 499 filter VALUES across all filters. The maximum number of cut-point tags is 99, with an overall maximum of 499 cut-point tag values. The maximum number of lump tags is 99, with an overall maximum of 499 lump tag values.

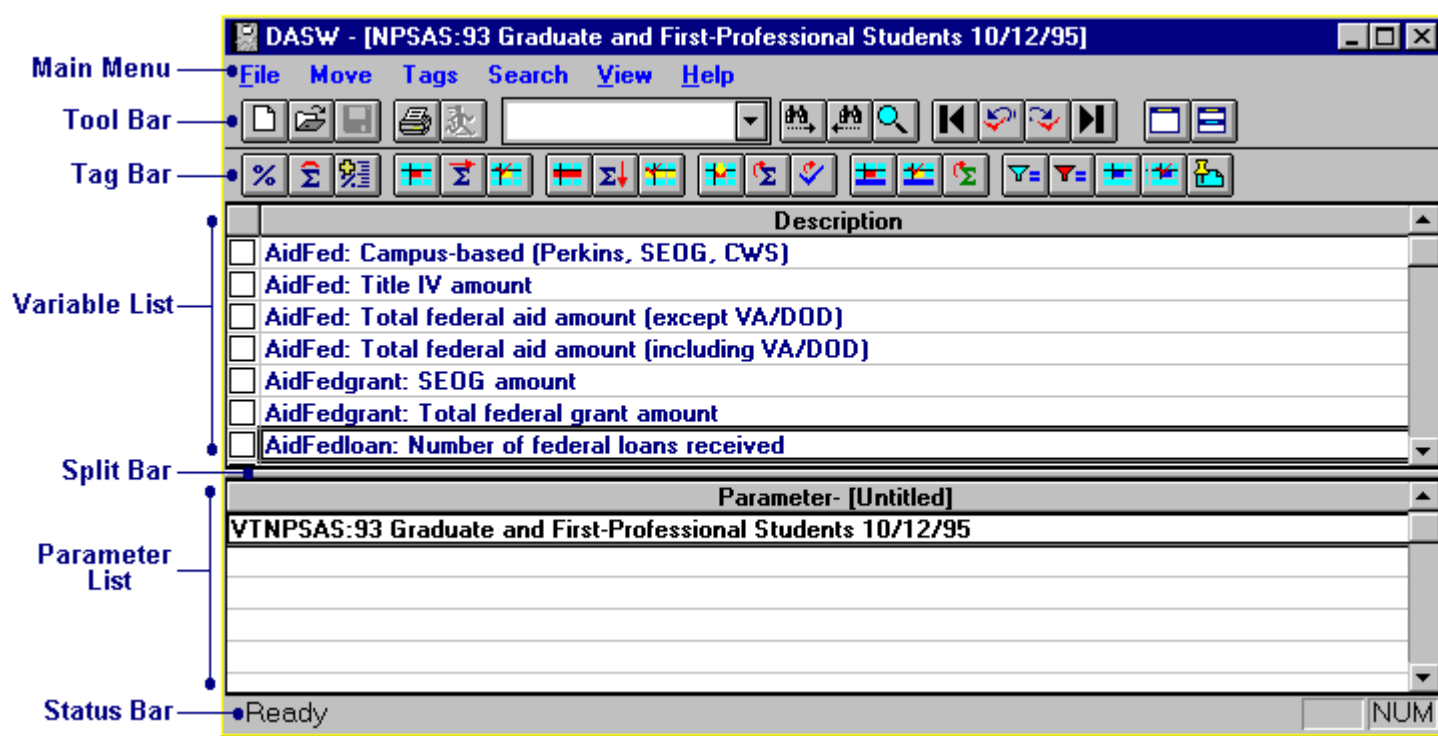
There is an overall maximum of 299 variables when running a correlation matrix.

The software that computes table estimates may be limited (beyond the above values) by the memory available in the user's PC. When problems are encountered, users should close all other programs and rerun the table or correlation matrix. If all else fails, try submitting the parameter file(s) over the Web; see <http://nces.ed.gov/DAS/>.

See also: System Requirements.

# The DAS Main Window

The DAS Main Window consists of the Variable List (Description), Parameter List menu, toolbar, tagbar and status bar.



## Variable List

The Variable List displays for each variable a **Section: Label description** and ☒ (check mark box) indicating whether the variable has been tagged for use. When you click the right mouse button in the Variable List, a context-sensitive menu for Tag commands (showing only the appropriate options for that variable) appears. You may use the mouse or keyboard (PgUp/PgDn) to navigate through the variable list. The color of the variable name and label indicates the variable type:

Color	Type
-----	-----
Green	Categorical
Navy Blue	Continuous
Dark Red	Weight

## Parameter List

The Parameter List displays the table specification, including its label. The table specification is created by tagging a variable. When you click on the right mouse button in the Parameter List, a menu appears with options including Edit and Delete.

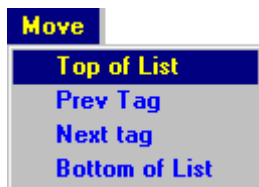
## File Menu

The File menu indicates the commands used to create parameter file, import parameter file, save file, edit parameter file, edit title, and change the system setup.



### **Move Menu**

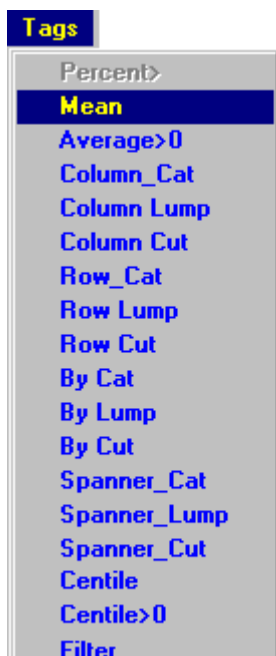
You can navigate through the variable list displayed in the main work space using commands from the Move menu.



### **Tag Menu**

Depending on the currently selected mode (table or correlation), there are two tag menus available. You create a parameter using commands from the Tag menu. You also bring up the Tag menu by clicking on the right mouse button, or choose an icon on the toolbar.

*Tag menu for tables:*



Tag menu for correlation matrices:



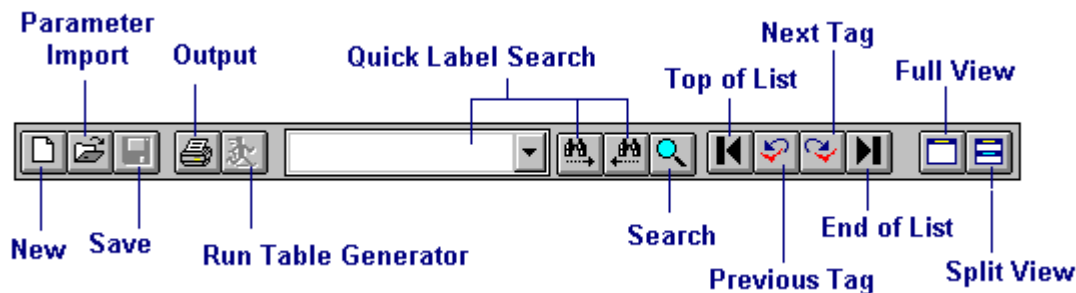
### View Menu

You can change the view mode of the screen, or hide the toolbar, tagbar, and status bar using commands from the View menu.



### Toolbar

The toolbar provides quick access to functions from the File, Move, Tag, and View Menu and to the Search Command.

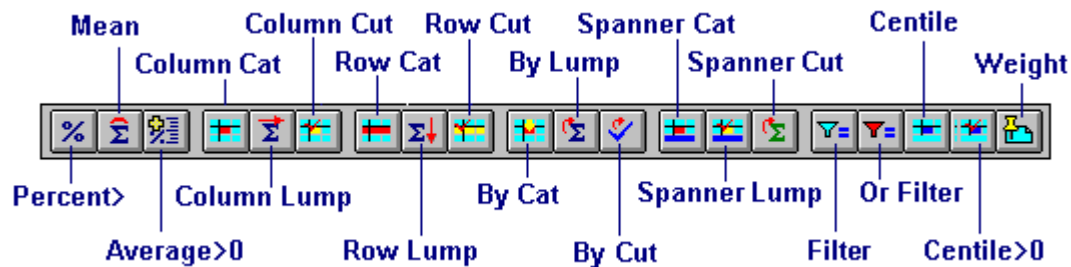


The Quick Label Search may be used to search on the **Section:** component of the variables' labels. After using a dataset for several tables, this quick label search is very useful.

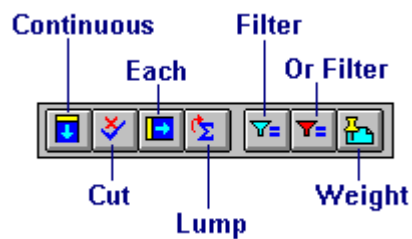
### Tagbar

The tagbar provides quick access to functions from the Tag menu. There are two types of tagbars, depending on the selected mode (i.e., table or correlation matrix).

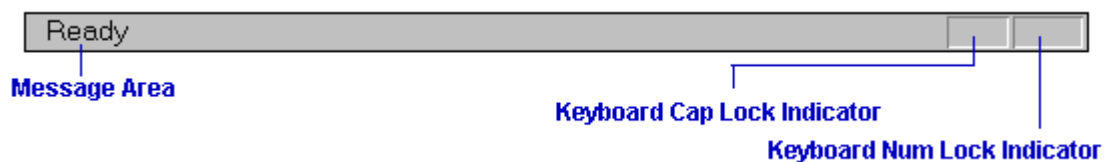
Tagbar for tables:



*Tagbar for correlation matrices:*



### **Status Bar**



When you select a command from a menu, a description of the menu item is displayed in the message area.

# How to View Description/Percentage

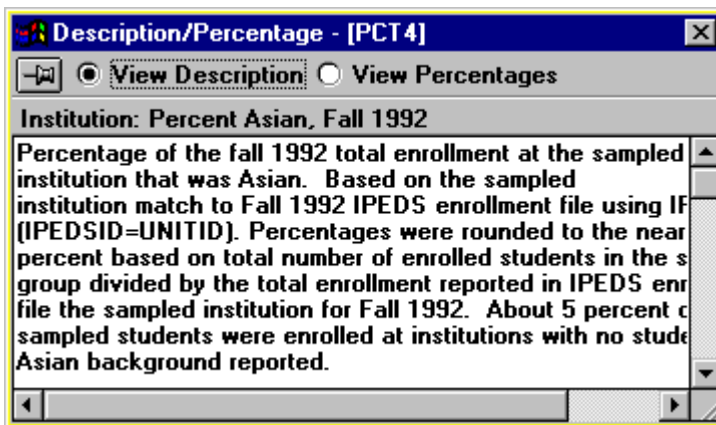
Percentage/Description Windows give descriptive/statistical information for the highlighted variable. NCES recommends strongly that users refer to these windows for important information about how variables were defined.

There are two ways to display variable descriptions:

1. Position the mouse arrow on the variable of interest and double click the left mouse button.

Or,

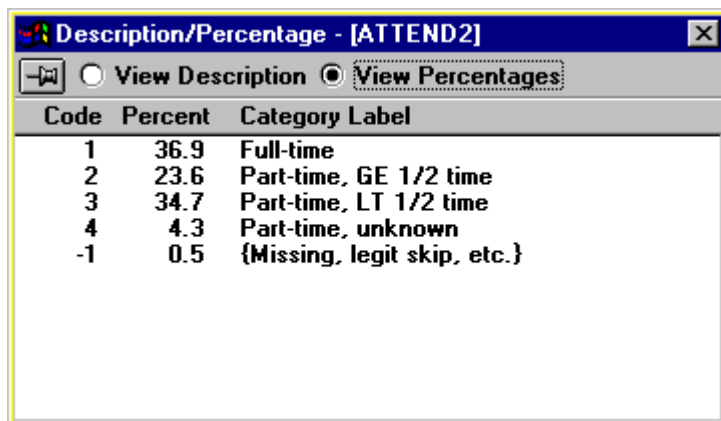
2. Use the up or down cursor key to position the black box over the variable entry, then choose the **Description** option from the **View** menu.



The Description window displays the descriptive information such as:

- Wording
- CATI/CAPI items
- Recoding specification
- Information on how computed if different from originally collected
- Notes about which files the data originated from, and if computed, by whom

You can view the percentage for the highlighted variable by clicking on the View Percentages button in the top of the window.







Code	Percent	Category Label
1	36.9	Full-time
2	23.6	Part-time, GE 1/2 time
3	34.7	Part-time, LT 1/2 time
4	4.3	Part-time, unknown
-1	0.5	{Missing, legit skip, etc.}

The percentage window displays the statistical information as follows:

- Values
- Labels
- Percentages(weighted)
- For continuous variables, minimum-maximum and Mean/Standard deviation are displayed within braces. The first value within the braces is the minimum value that is greater than zero, followed by the maximum value; then the mean followed by the standard deviation. The means and standard are rounded and based on weighted data. However, the standard deviations have not been adjusted for the survey design. This means that they are generally smaller than would be expected. When the DAS table produces the final estimates, the standard error will correctly take the design into account.

You can open up to two description windows simultaneously by repeating either of the above steps. You can move description windows using the standard Windows click-and-drag.

If you click on  button in the Description window, this button will be changed to . The  button prevents this window from displaying the description of another variable; it fixes the display in the window. If you choose to view the description of another variable, it will be displayed only in the description window with  button.



# How to Navigate Through Variables

You can move to the top of the variable list, to the bottom of the variable list, and to tagged items.

To move to the top of the list

Choose the Top List command from the Move menu.

Or, click on  button.

To move to the bottom of the list

Choose the Bottom List command from the Move menu.

Or, click on  button.

To move to the next tagged item

Choose the Next Tag command from the Tag menu.


Or, click on  button.

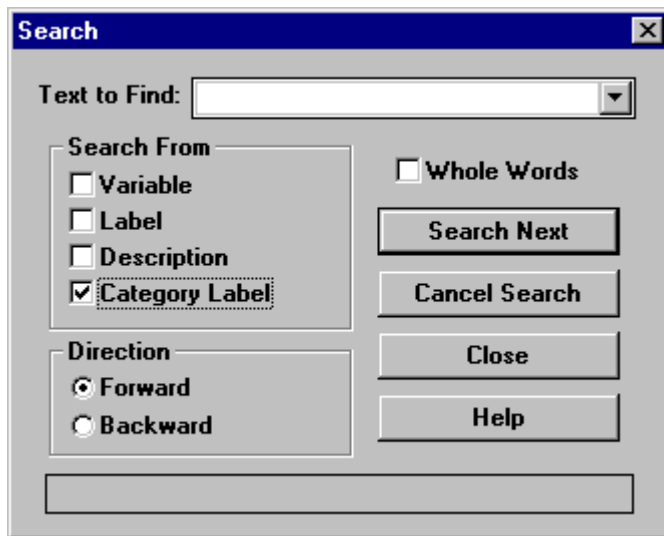
To move to the previous tagged item

Choose the Previous Tag command from the Tag menu.

Or, click on  button.

# How to Search for Text

Choose the search command from the main menu, or click on  button from the Toolbar to search for specified text.



## **Dialog Box Options**

### ***Text to Find***

Text to be found. You may enter text or click on the down arrow to the right and choose from the list displayed. The list displays the text string(s) previously entered, if any. All text comparisons are UPPER case, as are all text entries.

### ***Search From***

Selects the area(s) to be searched. The four possible areas are Variable name, variable Label, variable Description (as displayed in the description window), and Category Label (as displayed in the percentages window).

### ***Search Next***

Starts the search and stops at the first match. You must click Search Next to start the search; hitting enter after entering the Search Next does not start the search.

### ***Whole Words***

Finds occurrences that are whole words and not part of a larger word.

### ***Direction***

Selects the direction of the search—Forward toward the end of the list or Backward toward the beginning of the list. The search always starts from the current location of the cursor (black rectangle). Many users Move to Top of List before initiating any search.

### ***Cancel Search***

Cancels the search. When searching includes Descriptions, users may find the search much slower than they would like.

### ***Close***

Closes the search dialog box and returns back to previous action.

# How to Create a List of Parameters

In creating a list of parameters, you select variables for tables or correlations. This process requires making choices about how they are used. Designing effective tables involves decisions about placement (as row or column), use of filters to focus on specific cases, design of cut-points, aggregates (or lumps), and other variable recodes or transformations. These are complex tasks and all can be performed from the Tag menu.

There are three ways to create a parameter. You can:

1. Select a variable by using the arrow key or mouse, then choose a desired option from the Tag menu.

**Or,**

2. Select a variable by using the arrow key or mouse, then press the right button on the mouse. A context sensitive floating tag menu pops up right below the highlighted variable. You then choose an option from the Tag menu.

**Or,**

3. Select a variable by using the arrow key or mouse, then click on the button in the tag bar below the main menu.

The Tag menu provides a list of possible variable tags. The most frequently used are Filter, Row Cat, Column Cat, Mean, Average>0, and Percent>. The Tag menu also defines the Title(s) for the table. This feature allows users to specify a different set of value labels as rows in the table, such as income categories or financial aid amount groupings. When a tag item has been selected, a corresponding pop-up window is displayed. You then fill in the columns on the corresponding window. When you save the data entered by clicking on "OK" button, ☒ appears on the left corner of the highlighted variable, and the data saved shows up on the last row of the Parameter List. Variables may have more than one tag. Identical tags on the same variable are allowed ONLY for Row Cut and Filter tags.

To find out details on the corresponding popup window, click on the tag below.

For the table,

**Percent>**  
**Mean**  
**Average>0**  
**Centile**  
**Centile>0**  
**Column\_Cat**  
**Column\_Lump**  
**Column\_Cut**  
**Row\_Cat**  
**Row\_Lump**  
**Row\_Cut**  
**By\_Cat**  
**By\_Lump**  
**By\_Cut**  
**Spanner\_Cat**

**Spanner\_Lump**  
**Spanner\_Cut**  
**Filter**  
**Or\_Filter**  
**Weight**



For the correlation matrix,

**Continuous**  
**Cut**  
**Each**  
**Lump**  
**Filter**  
**Or\_Filter**


See also: [How to Delete a Parameter](#), [How to Edit a Parameter](#).

# How to Resize Parameter List

There are three ways to change the size of the Parameter List.

1. Click on  button to minimize it, or click on  button to make it half the size of the screen.
2. Select the *Full* option from the *View* menu to minimize it.

**Or,**

- 3.a. Select the *Split* option from the *View* menu to make it half the size of the screen.
- 3.b. Point the mouse to the Split bar. When the pointer changes to , drag it down or up to the position you want.

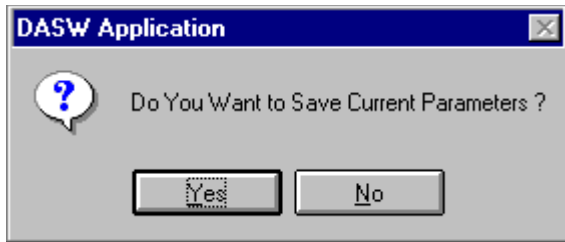
# How to Create a New Parameter File

1. Choose the *New* option from the File menu.

**Or**

2. Click on  button.

3. A popup window appears asking you for the confirmation to save the parameters entered.



4. When you click on the "Yes" button, the currently defined parameters on the Parameter List are saved into a file.

5. When you click on the "No" button, the currently defined parameters on the Parameter List are discarded.

6. The Parameter List is cleaned except the first line which is the title of the project.

7. The Parameter List displays "Untitled" as a file title.

8. Select variables for a new table or import a parameter list.


See also: [How to Save a Parameter File.](#)

# How to Edit a Parameter

When a table's parameters are defined for rows, columns, and filters based on categorical data, the current labels for the input values are displayed on the Parameter List.

You can edit a parameter by performing the following steps:

Select a row from the Parameter List by positioning the mouse pointer and clicking on the row you want to edit.

Press the right button on your mouse. Then a floating menu  appears right below the row highlighted.

Choose the **Edit** option from the menu.

**Or,**

Double-click on the row selected from the Parameter List.

Then an appropriate popup window appears. To find out details on the popup window, click on the tag name below.

For the table parameters:

- Percent>**
- Mean**
- Average>0**
- Centile**
- Centile>0**
- Column\_Cat**
- Column Lump**
- Column Cut**
- Row\_Cat**
- Row\_Lump**
- Row\_Cut**
- By\_Cat**
- By\_Lump**
- By\_Cut**
- Spanner\_Cat**
- Spanner\_Lump**
- Spanner\_Cut**
- Filter**
- Or\_Filter**
- Weight**

For the correlation matrix parameters:

- Continuous**
- Cut**
- Each**
- Lump**

**Filter**  
**Or\_Filter**

See also: [How to Delete a Parameter](#), [How to Create a List of Parameters](#).




# How to Delete a Parameter

When parameters are defined, the current labels for the input values are displayed in the Parameter List.

You can delete a parameter by performing the following steps:

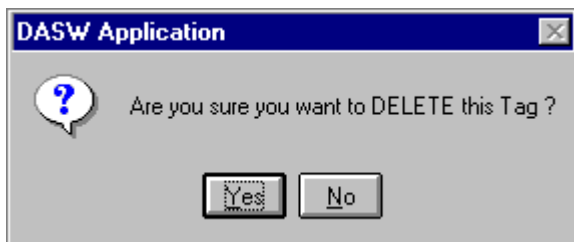
1. Select a row from the Parameter List by positioning the mouse pointer and clicking on the row to be deleted.

Or,

2. Press the right button on the mouse. Then a floating menu  pops up right below the row selected.


3. Choose the **Delete** option from the menu.

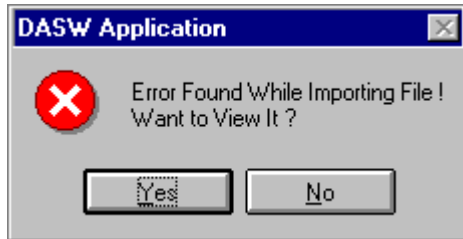
4. Then a popup window appears asking you to confirm the deletion of the selected row.



5. When you click the "OK" button, the highlighted parameter is removed from the Parameter List.

# How to Import a Parameter File

You can retrieve and edit old parameter files previously saved by choosing the **Import** option from the **File** menu, or by clicking on  button. When DASW imports the file, it performs thorough error checking. When errors are detected, an error log file (ERROR.TXT) is created. A popup dialog then appears, prompting the user to confirm.



When you click on the **Yes** button, the error log will be displayed using your Parameter File Editor (usually NOTEPAD).

*WARNING: DASW will import very OLD.tpf files that may use old versions of variables. If you create a table and encounter a large number of rows filled with "Low n" estimates after importing an old .tpf file, check the values specified by comparing with the percentage window information.*

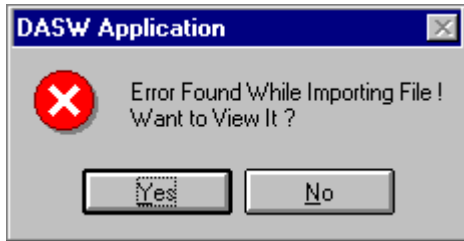
# How to Edit a Parameter File

Choose the **Edit Parameter File** option from the **File** menu.

This option allows you to edit a parameter file currently displayed on the parameter list window. You can change the editor used by changing the File Setup.

*NOTE: Make sure to save the parameter file after you complete editing using the **Save** option from the file editor.*

When you close the **File Edit** window, DASW imports the file while performing extensive error checking. When errors are detected, an error log file (ERROR.TXT) is created. A popup dialog then appears, prompting the confirmation for viewing the error log.

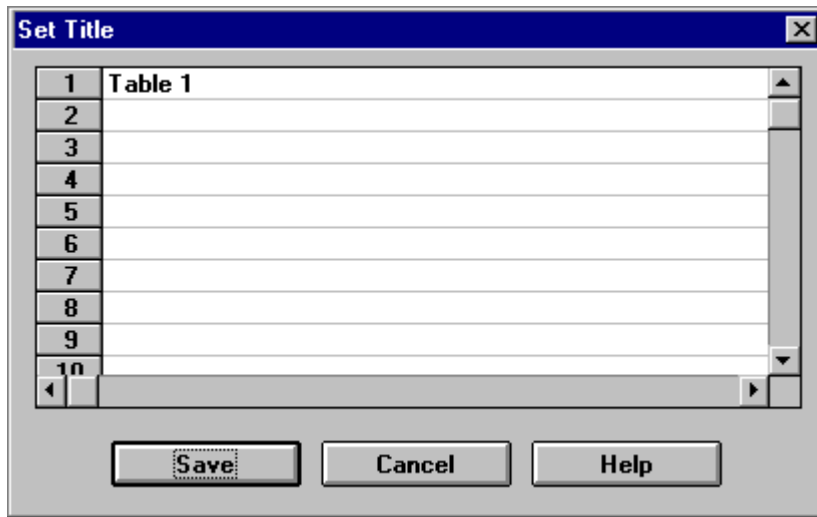


When you click on **Yes** button, the error log will be displayed on Windows's Notepad application.

See also: [How to Change the System Setup](#), [How to Import Parameter File](#) .

# How to Edit Titles

Choose the **Edit Title** option from the **File** menu. The title dialog window appears as follows:



The image shows a dialog box titled "Set Title" with a close button (X) in the top right corner. Inside the dialog, there is a table with 10 rows. The first row is labeled "1" and contains the text "Table 1". The remaining rows are labeled 2 through 10 and are empty. To the left of the table is a vertical scrollbar, and to the right is a horizontal scrollbar. At the bottom of the dialog, there are three buttons: "Save", "Cancel", and "Help". The "Save" button is highlighted with a dashed border.

1	Table 1
2	
3	
4	
5	
6	
7	
8	
9	
10	

This option allows you to specify titles for the table. You can enter up to 50 lines. Click on the Save button to save the lines entered.

# How to Save a Parameter File

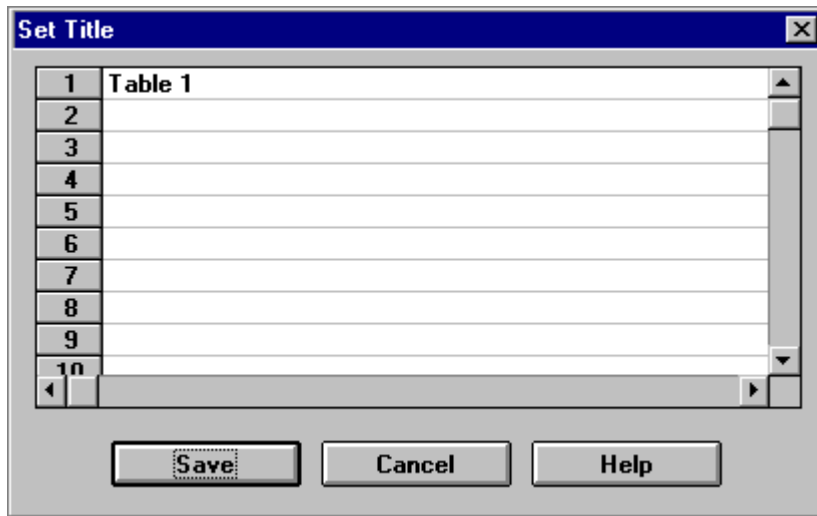
1. To update an old parameter file, select a valid parameter file from the display list, then click on the **Save** option from the **File** menu to resave it under its original name or **Save As** option from the **File** menu to give it a new name.

Or,

2. Click on  button.

3. If this is a new file, a standard Windows directory dialog box appears, permitting you to specify the directory and file name for the output file.

4. The title dialog window pops up waiting for you to enter the title.



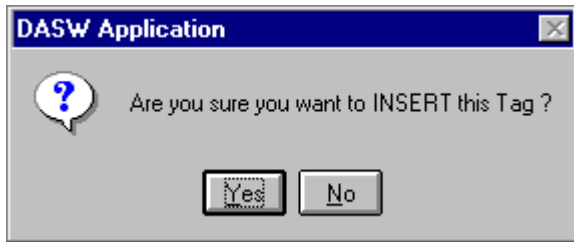
The default title is a list of variables you have tagged.

5. When you click on the **Save** button, the Parameter List will be saved into a file stored in the output location.

See also: Setup.

# How to Insert a Row

1. Click the right mouse button on the row above which you wish to insert a row.
2. Then a floating menu appears.
3. Choose the Insert option from the menu.
4. A dialog window will prompt a confirmation for a new row.



5. Click on the **Yes** button to insert a new row. The new row will be above the row clicked in #1.

# How to Delete a Row

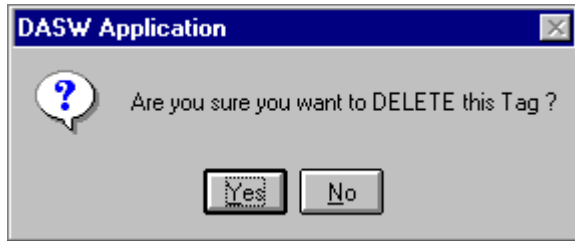
1. Click the right mouse button on the row you wish to delete.

2. Then the context-sensitive menu appears.

*Note: If you are not allowed to delete the currently chosen row, the menu will NOT appear.*

3. Choose the Delete option from the menu.

4. A dialog window will prompt a confirmation.




5. Click on the **Yes** button to delete the selected row.

# How to Get Tables

There are two ways to process your .tpf files. If you have a CD-ROM version of the DAS, please see Running Tables On Your Computer. If you downloaded the DAS software from [www.nces.ed.gov](http://www.nces.ed.gov), please see Running Tables Via DAS FTP Site.

## **Running Tables On Your Computer**

After building (and saving) a parameter file, choose the Run Table option  from the main menu.

If you have already created and saved several .tpf files, you may run them all in batch mode. To run files in batch mode, go to a DOS prompt and then change (CD\ ) to the directory containing your DASW program. From there, enter the following command:

```
DAST p=<pid> <path>
```

where <pid> is the 3-digit project identification code and <path> is the location of the .tpf files.

For example, the command

```
C:\DASW\DAST p=N6U c:\dasw\n6u\*.tpf
```

would run all .tpf files in c:\dasw\n6u. The output will be saved with the same name as the .tpf file but with the extension .prn.

## **Running Tables via DAS FTP site**

After building (and saving) parameter files, compress your .tpf files into a .zip and submit that file to the DAS FTP site to get tables. The software (programs) you will need and the corresponding operation instructions are available through [www.nces.ed.gov/das/htm/submit.htm](http://www.nces.ed.gov/das/htm/submit.htm).




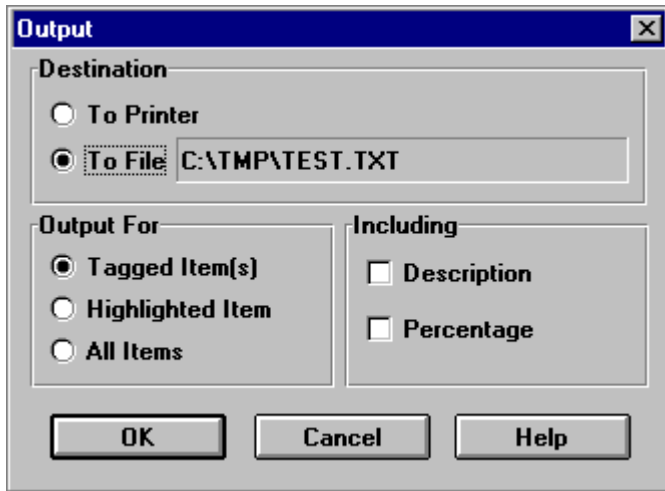
## How to View DAS Output

Tables, DAS output files, will be either have the extension .prn or .csv. Other than the obvious difference in file names, the only difference between .csv and .prn files is the column heading string. In .csv files, the column heading are written as one long string. In .prn files, the column headings are broken into several rows of strings (the break point character is a blank). Depending on your spreadsheet program, you may want to change the file extension generated by the DAS. You should use .csv files for Excel and .prn files for Lotus or Quattro. The DAS default extension is .prn. Instructions for selecting the file extension generated by the DAS tables are available in [How to Change the Systems Setup](#).

Alternatively, you can view your tables using PRNFORM, the table output file viewer provided with the DAS application. PRNFORM is a simple program which can turn a .prn or .csv file into two different text files, which can be easily read with a text editor. The files will have the same name as the original .prn or .csv file, but with a .txt or .se extension. The .txt file will contain the percentages or mean distributions for the tables, while the .se file will contain the standard errors directly under the percentages or means distributions. Instructions on how to use PRNFORM are provide with the DAS application (prnform.txt) and also available at [http://nces.ed.gov/das/htm/download\\_prnform.htm](http://nces.ed.gov/das/htm/download_prnform.htm).

# How to Print Variable Information

Choose the **Output** option from the main menu or click on  icon in the toolbar to get printed information about variables. The following dialog box will appear:



## ***Destination***

The output mode by which the codebook is generated. To print the output, choose **To Printer**. Or choose **To File** to save the output into a file--the Windows standard directory dialog box that pops up allows you to choose the file name. Click on the **OK** button to return to the output window. You can change the file name only by choosing **To File**.

## ***Output For***

The range of the variables included in the output.

## ***Including***

To include the description in the output, choose **Description**. To include the percentage, choose **Percentage**. If nothing is chosen, only the variable list is included in the output.

## ***OK***

Prints or saves the output, then returns to the main menu.

## ***Cancel***

Cancels the data entered and returns to the main menu.

# How Do I ?

## ***Find a variable?***

Use the up and down cursor to scroll through the variables list (this is the slow way).

Use the SEARCH function:

If you know the variable name, enter it in the text box and click "search".

Otherwise, enter words in the text box that you expect to see in the variable label or description windows.

Use the prefix search, if you are familiar with the variables.

## ***Choose between two similar variables?***

Read both VIEW DESCRIPTION windows and decide.

## ***Pick the correct weight?***

Read both VIEW DESCRIPTION windows for each weight and pick the weight best for your dependent variable (columns).

## ***Define ranges from a continuous variable into a table as rows?***

Use a Row\_Cut.

## ***Determine the distribution of a continuous variable in a table?***

Read the VIEW PERCENTAGES window.

Use a Column\_Cut.

Use a Centile.

Use a Centile>0.

## ***Define the subsample for estimates based on a range of a continuous variable?***

Use two Filters for the variable. For example, to define the subsample with INCOMES between \$20,000 and \$30,000, use the following FILTER parameters:

```
FILTER INCOME Adjusted gross income  
19999.5 $20,000 or more  
FILTER INCOME Adjusted gross income  
-30,000.5 $30,000 or less
```

## ***Combine several categories for a variable together as a new category?***

Use a Lump.

## ***Get the grand totals when using a SPANNER tag ?***

This cannot be done with a .tpf file that includes the SPANNER. You can import that .tpf file, delete the SPANNER, and run this smaller .tpf file to get totals.

## ***Make the table system work with my spreadsheet software?***

Use FILE SETUP: For Lotus, use .prn and for Excel use .csv. Browse to find your spreadsheet program (EXCEL.EXE, 123.BAT, etc.).

## ***Make the DAS create correlations?***

Read the HELP file on HOW TO START DASW.

***Adjust t-tests in regressions for design effects?***

Divide the calculated  $t$  from SPSS or SAS by the design effect computed for the dependent variable. The design effects are the DEFTs displayed at the bottom of RMATRIX.PRN.

***Get a printout of all the variables in a DAS system?***

Read the HELP file on How to Generate Printed Output

***Get a frequency distribution for a variable?***

See VIEW PERCENTAGES window; in a DAS table, use a Column\_Cut, Centile or Centile>0.

***Create a new variable in the DAS?***

This option is NOT available. Sums of estimates will sometimes work. Call NCES (Aurora D'Amico 202.219.1365) and ask that the new variable be added. Please describe the new variable with SAS or SPSS code.

***Get a median for a variable?***

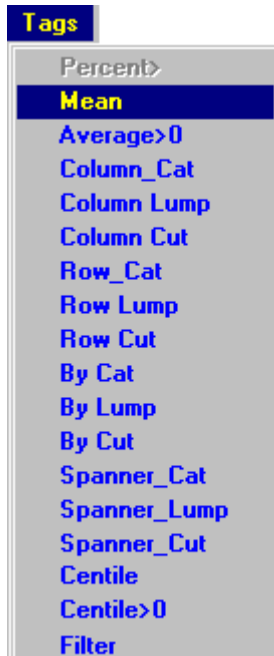
Use a Centile.

Use a Centile>0.

# All About Tags

Depending on the currently selected mode (table or correlation), there are two tag menus available. You can create a parameter using commands from the Tag menu. You can also bring up the Tag menu by clicking on the right mouse button, or choose an icon on the toolbar.

*Tag menu for the table:*



*Tag menu for the correlation matrix:*



See also:

For the table,

**Percent>**

**Mean**

**Average>0**

**Centile**

**Centile>0**

**Column\_Cat**

**Column\_Lump**

**Column\_Cut**

**Row\_Cat**

**Row\_Lump**

Row\_Cut  
By\_Cat  
By\_Lump  
By\_Cut  
Spanner\_Cat  
Spanner\_Lump  
Spanner\_Cut  
Filter  
Or\_Filter  
Weight

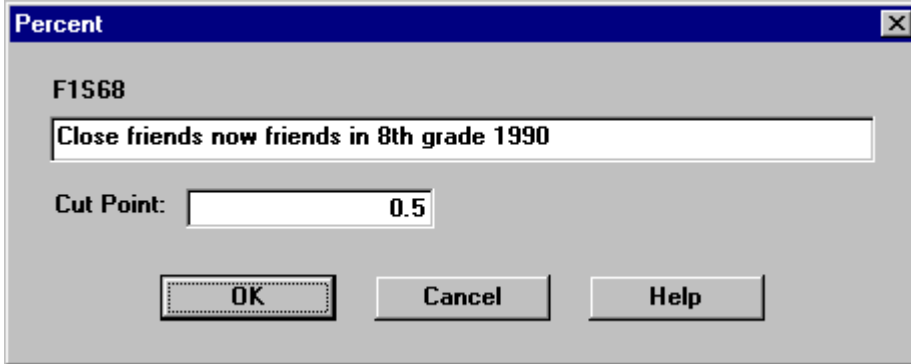
For the correlation matrix,

**Continuous**  
Cut  
Each  
Lump  
Filter  
Or\_Filter

## Percent>

Percent> builds a column of percentages where the numerator is the weighted number of cases greater than the cut-point value and the denominator is the weighted number of cases greater than -1. When used with a By\_Cut, By\_Cat or By\_Lump tag, only **one** Percent> tag may be specified. Otherwise, several Percent> tags may be defined in a single table. Also, Percent> tags may be included with Mean and Average>0 tags. Percent> tags may NOT be mixed with Column\_Cut, Column\_Cat, Column\_Lump, Centile, or Centile>0 tags.

When you create or edit a Percent> tag, the following dialog box appears:

A screenshot of a dialog box titled "Percent". The dialog box has a blue title bar with a close button (X) in the top right corner. The main area is light gray. At the top, the text "F1S68" is displayed. Below it is a text input field containing the label "Close friends now friends in 8th grade 1990". Underneath the text field is a label "Cut Point:" followed by a numeric input field containing the value "0.5". At the bottom of the dialog box are three buttons: "OK", "Cancel", and "Help".

Users should edit the label and enter the Cut Point. When you click on the OK button, the information is saved into the Parameter List. **The default cut-point value is 0.5 with the existing label.**

### Examples:

To specify the percentage of students with family incomes of \$20,000 or more, you would use:

```
Percent>19999.5 INCOME Adjusted gross family income
```

To specify the percentage receiving Pell grants, you would use:

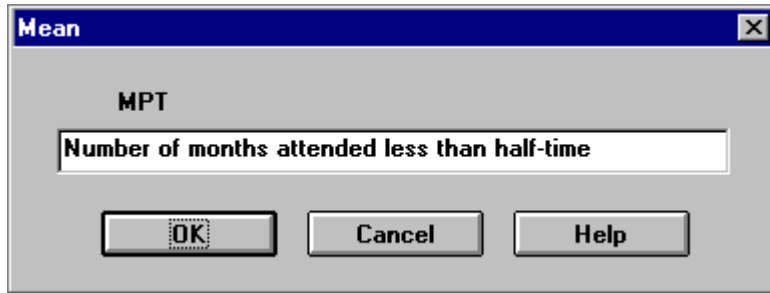
```
Percent>0.5 PELLAMT Pell grant amount
```

See also: Average>0, Mean, Column\_Cut, Column\_Cat, Column\_Lump, Centile, or Centile>0.

# Mean

Mean defines a column of arithmetic means in the table. Negative (-1) values are treated as missing; however, zero is valid. When used with a By\_Cut, By\_Cat or By\_Lump tag, only **one** Mean tag may be specified. Otherwise, several Mean tags may be defined in a single table. Also, Mean tags may be included with Percent> and Average>0 tags. Mean tags may NOT be mixed with Column\_Cut, Column\_Cat, Column\_Lump, Centile, or Centile>0 tags.

When you create or edit a Mean tag, the following dialog box appears:



Edit the label and click on the OK button to save the information into the Parameter List.

## Example:

To specify the mean PELL grant, including zero amounts for non-recipients, you would use:

```
Mean PELLAMT PELL grant amount
```

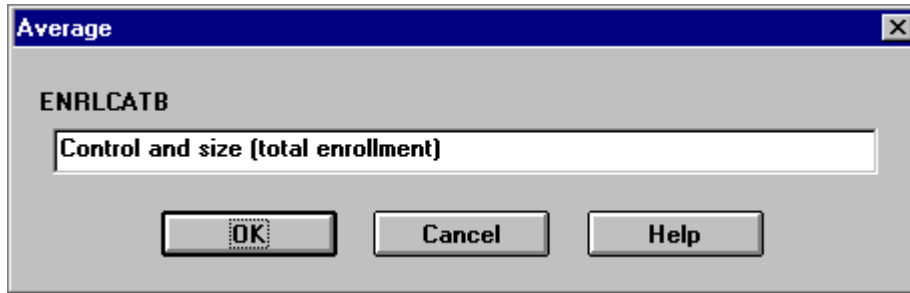
See also: Average>0, Percent>, Column\_Cut, Column\_Cat or Column\_Lump, Centile, or Centile>0.



## Average>0

Average>0 defines a column of averages in the table. Average includes only positive values (zeros and missing values are not included in these calculations). When used with a By\_Cut, By\_Cat or By\_Lump tag, only **one** Average>0 tag may be specified. Otherwise, several Average>0 tags may be defined in a single table. Also, Average>0 tags may be included with Mean and Percent> tags. Average>0 tags may NOT be mixed with Column\_Cut, Column\_Cat, Column\_Lump, Centile, or Centile>0 tags.

When you create or edit an Average>0 tag, the following dialog box appears:



Edit the label and click on the OK button to save the information into the Parameter List.

### Example:

To specify the average PELL grant received (excluding zero values for non-recipients), you would use:

```
Average PELLAMT PELL grant amount
```

See also: Mean, Percent>, Column\_Cut, Column\_Cat or Column\_Lump, Centile, or Centile>0.

# Centile

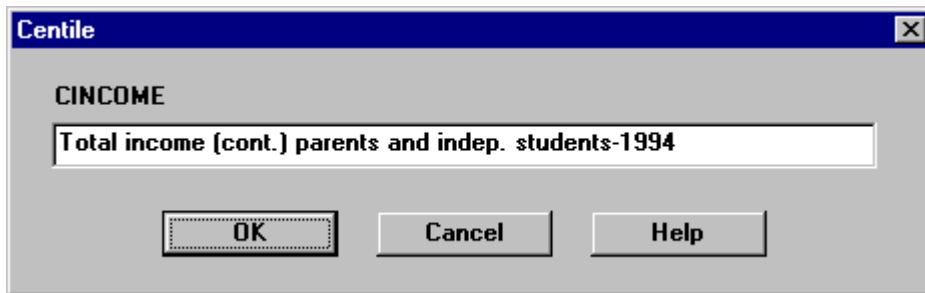
Centile generates percentile columns in your table from continuous variables. Values for the following five percentiles are produced: 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup> (median), 75<sup>th</sup>, and 90<sup>th</sup>. The 10<sup>th</sup> percentile value represents the lowest decile while the 90<sup>th</sup> percentile value is the highest decile.

During tag creation, it is not possible to specify a subset of the percentiles, although after creating the table, you can delete undesirable columns.

Centile can only be used as a column variable and may not be used in combination with any other column tag. Minimum and maximum values are automatically entered.

## Dialog box:

Edit the When you create or edit a Centile tag, the following dialog box appears.



Edit the label button and click on the OK button to save information into the Parameter List.

## Examples:

To determine percentile CINCOME Total income, the parameter would contain:

```
Centile CINCOME Total income (cont.) parents and indep. students-1994
C 1.00 2000000.00
```

See also: Centile>0.

## Centile>0

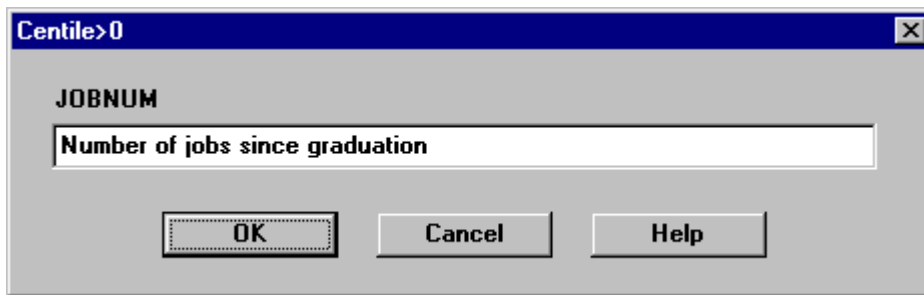
Centile>0 generates percentile columns in your table from continuous variables. This function filters out all zero values and starts with the first, valid non-zero value. Values for the following five percentiles are produced: 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, (median), 75<sup>th</sup>, and 90<sup>th</sup>. The 10<sup>th</sup> percentile value represents the lowest decile, the 90th percentile value is the highest decile. In addition, a sixth column is generated showing the percentage of cases with a zero value. If your desired variable does not have zero values, Centile>0 will not generate any values for the 0<sup>th</sup> percentile and will be identical to the output generated by the Centile tag.

During tag creation, it is not possible to specify a subset of the percentiles, although after creating the table, you can delete undesirable columns.

Centile>0 can only be used as a column variable and may not be used in combination with any other column tag. Minimum and maximum values are automatically entered.

### Dialog box:

When you create or edit a Centile>0 tag, the following dialog box appears.



Edit the label and click on the OK button to save the information into the Parameter List.

### Examples:

To determine percentiles for number of jobs since graduation, where the minimum value is zero and maximum value is eleven, you would use:

```
Centile>0 JOBNUM Number of jobs since graduation  
C 0.00 11.00
```

See also: Centile.

# Column\_Cat

Column\_Cat uses **ALL** of a categorical variable's categories to form columns of percentages (that sum to 100). You may specify only ONE Column\_Cat variable.

The dialog box appears as follows:

	Code	Label
<input type="checkbox"/>	1	Public, under 1,000 enrolled
<input type="checkbox"/>	2	Public, 1,000-2,499
<input type="checkbox"/>	3	Public, 2,500-4,999
<input type="checkbox"/>	4	Public, 5,000-9,999
<input type="checkbox"/>	5	Public, 10,000-19,999
<input type="checkbox"/>	6	Public, 20,000 or more
<input type="checkbox"/>	7	Private, non-profit, under 1,000
<input type="checkbox"/>	8	Private, non-profit, 1,000-2,499
<input type="checkbox"/>	9	Private, non-profit, 2,500-4,999
<input type="checkbox"/>	10	Private, non-profit, 5,000-9,999
<input type="checkbox"/>	11	Private, non-profit, 10,000 or more
<input type="checkbox"/>	12	Private, for-profit, under 300

Users may edit the variable label and any value labels. When you are satisfied with the labels, click on the OK button.

See also: Column\_Lump.

# Column\_Lump

Column\_Lump produces columns summing across categories to 100 percent in the table. You may use only one Column\_Lump per table. Lumping operates on categorical variables to define new groupings of categories.

The dialog boxes for the entry of the lump specifications are the same as for a Row\_Lump.

See also: Column\_Cat, Row\_Lump.

# Column\_Cut

Column\_Cut (for columns of percentages) may be used to group contiguous categories or break up a continuous variable into categories to form columns in your table. Only one Column\_Cut may be specified per table.

Column\_Cut variables are like Row\_Cut, only they form columns of percentages for continuous variables or for new contiguous groupings of any ordered variable. The percentages estimated for a Column\_Cut will sum to 100 percent. The dialog boxes for the entry of values is the same as for ROW\_CUTs .

See also: Column\_Cat, Column Lump, Row\_Cut.

# Row\_Cat

Row\_Cat uses **ALL** the variable's categories to form rows in the table. Row\_Cat operates only on categorical variables. You may define several Row\_Cat variables, but Row\_Cat may only be used once for a single categorical variable.

For categorical variables that do not need any transformations (e.g., lumping), Row\_Cat tags specify use of all the variable's categories as rows in the table. That is, each category becomes a row defining the denominator of the estimates. You can use Row\_Lump, however, to select or limit the categories shown as rows. For example, if you wanted to show all categories for institution type, and also produce estimates by combining all 4-year institutions, you would tag the variable for institution type as a Row\_Cat, then tag again as Row\_Lump.

The dialog box for editing the labels for a Row\_Cat tag appears as follows:

	Code	Label
1	1	Public, under 1,000 enrolled
2	2	Public, 1,000-2,499
3	3	Public, 2,500-4,999
4	4	Public, 5,000-9,999
5	5	Public, 10,000-19,999
6	6	Public, 20,000 or more
7	7	Private, non-profit, under 1,000
8	8	Private, non-profit, 1,000-2,499
9	9	Private, non-profit, 2,500-4,999
10	10	Private, non-profit, 5,000-9,999
11	11	Private, non-profit, 10,000 or more
12	12	Private, for-profit, under 300

You may edit the variable label and the value labels as desired. When you are satisfied with the labels, click on the OK button.

See also: Column\_Cat, By\_Cat, Row\_Lump, and Spanner\_Cat.

# Row\_Lump

Row\_Lump produces rows of estimates in the table. Lumping operates on categorical variables to define new groupings of categories. For example, using a race/ethnicity variable with categories (1=Black, 2=White, 3=Hispanic, 4=American Indian), lumping categories 1+3+4 into a new category labeled "Minority" is frequently useful.

The first of two dialog boxes appears as follows:

	Code	Title
1		
2		
3		
4		
5		
6		

## To Enter Data

First, click on the **Add Lump** button. The second dialog box will appear as follows:



**Row\_Lump** [X]

F3RACE Race/ ethnicity 1994

	Code	Title
1		
2		
3		
4		
5		
6		

**Enter Lump Row**

**Categories:** Missing, legit skip, etc.  
Asian or Pacific Islander  
Hispanic regardless of race  
Black not of Hispanic origin  
White not of Hispanic origin  
American Indian or Alaskan native

**Lump Title:**

Second, select the categories desired for the grouping or lump from within the **Categories:** box. Third, enter a label in the **Lump Title:** box (if you selected only one category, its label will be the default **Lump Title:** label). Continue defining lumps as described in the second step by clicking on the **Add More** button.

Third, when you are finished defining lumps, click on the **Save & Close** button. This will return you to the first dialog box.

Fourth, examine the labels. If you need to edit the labels, click on the **Edit Lump** button.

Fifth, when you are satisfied with the categories that have been grouped together and the labels, click on the **Save & Close** button.

If you mistakenly click on **Add More** when you are finished creating your lumped categories, click on **cancel** and then **Save & Close**.

See also: Row\_Cat, Column\_Lump, and By\_Lump.

# Row\_Cut

Row\_Cut defines rows in your table from continuous variables or from contiguous ranges within a categorical variable. Cutting converts continuous values into a set of user defined categories. The Cut-Point values must be fractional (e.g., 99.5) and begin with the smallest value in the definition of a range.

You may use a variable more than once as a Row\_Cut.

## Dialog box:

After tagging the variable, a dialog box will appear for editing labels and entering Cut-Point values.

	Cut Point	Label
1	0.5	Less than Full-time full-year
2	7.5	8 or more months full-time
3		
4		
5		
6		
7		
8		
9		
10		
11		

First, edit the initial variable label--the new set of categories rarely will use the same label as the continuous variable. Second, enter the smallest cut-point value (0.5 above) and the label for the first range of values (e.g. Less than full-time full-year). Third, enter the next cut-point (e.g., 7.5) and the corresponding label (e.g., 8 or more months full time). Continue entering values and labels until the last range is defined. Note that the upper cut-point for the last range is not entered (a value of infinity is used). Finally, carefully scan the labels and edit as needed. When you are satisfied, click on the OK button.

## Examples:

Cut-point values must be fractional (e.g., 999.5) so all values will be either greater or less than the value. Cut-point values begin with the low end of the range. For example, to break up the range 0 through 2,000, cut-points might be:

```
Cut-point  Label
-0.5 Below 400      /* in this example, 0 is included */
399.5 400 to 699
699.5 700 or more
```

```
Cut-point  Label
-1.5 Missing      /*in this example, missing is included*/
```

```
-0.5  0-399      /* in this example, 0 is included */
399.5 400 to 999
999.5 1,000 or more
```

Cut-point Label

```
-1.5  Missing    /*in this example, missing is included*/
-0.5  Zero       /*this row includes only 0*/
0.5   1-399      /*includes greater than 0, AND < 400*/
399.5 400 to 999
999.5 1,000 or more
```

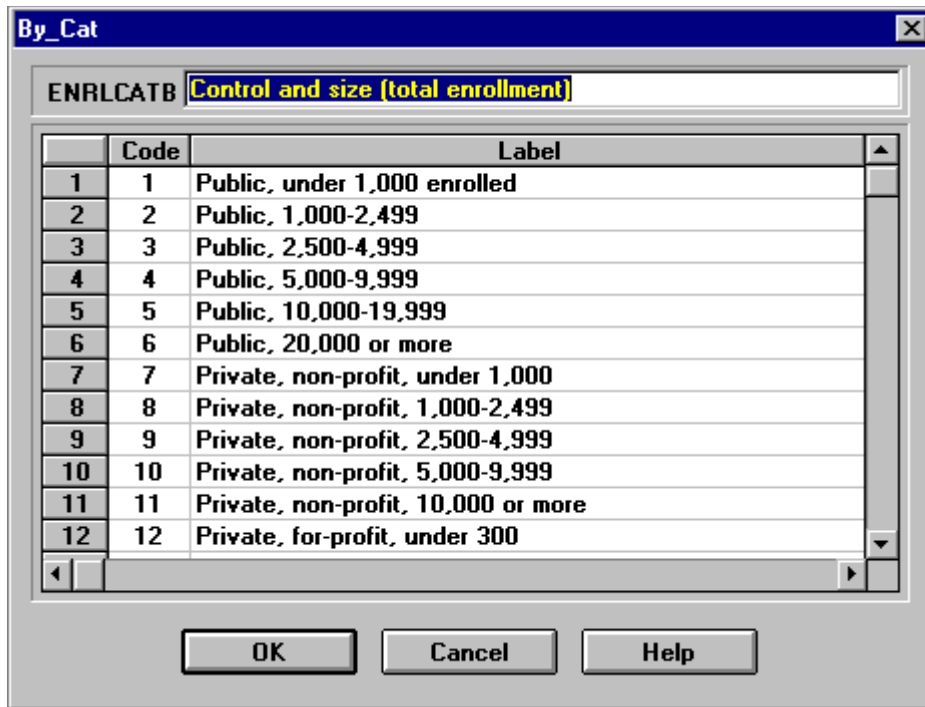
**NOTE: THIS VERSION OF THE DATA ANALYSIS SYSTEM ALLOWS CUT-POINTS THAT ARE NEGATIVE. THIS MAY BE HELPFUL IN EXAMINING MISSING VALUES WHICH ARE CODED AS -1. THIS ALLOWS USERS TO INCLUDE MISSING VALUES AS A SEPARATE ROW IN THE TABLE.**

See also: Row\_Cat, Column\_Cut, By\_Cut , Spanner\_Cut.

# By\_Cat

By\_Cat variables use **all** of a variable's categories to group columns of averages, means or percentage positives. BY tags may be used with only ONE Mean, Average>0, or Percent> variable.

When you create or edit a By\_Cat tag, the dialog box appears so that you can update the labels.



The dialog box titled "By\_Cat" has a text field labeled "ENRLCATB" containing the text "Control and size (total enrollment)". Below this is a table with 12 rows. The first two columns are "Code" and "Label". The "Code" column contains numbers 1 through 12. The "Label" column contains descriptions of enrollment categories. At the bottom of the dialog are three buttons: "OK", "Cancel", and "Help".

	Code	Label
1	1	Public, under 1,000 enrolled
2	2	Public, 1,000-2,499
3	3	Public, 2,500-4,999
4	4	Public, 5,000-9,999
5	5	Public, 10,000-19,999
6	6	Public, 20,000 or more
7	7	Private, non-profit, under 1,000
8	8	Private, non-profit, 1,000-2,499
9	9	Private, non-profit, 2,500-4,999
10	10	Private, non-profit, 5,000-9,999
11	11	Private, non-profit, 10,000 or more
12	12	Private, for-profit, under 300

Users may edit or replace the variable label and any value labels. When you are satisfied with the labels, click on the OK button.

**WARNING:** BY tags may dramatically increase the number of columns in a table.

See also: By\_Lump, By\_Cut.

## By\_Lump

By\_Lump operates with Mean, Average>0, or Percent> to produce columns of estimates in the table. Lumping operates on categorical variables to define new groupings of categories.

The dialog boxes for the entry of the lump specifications are the same as for a Row\_Lump.

See also: Column\_Lump, Row\_Lump, By\_Cat, Mean, Average>0, Percent>.

# By\_Cut

By\_Cut variables define groupings of columns from continuous variables or to group contiguous values within a categorical variable with one specified Average>0, Mean, or Percent> tag. Your table will include a column of means for each cut range specified and a total value.

**Only one** By\_Cut may be specified per table. **Only one** BY variable (either By\_Cat, By\_Lump, or By\_Cut) may be specified with **one and only one** Mean, Average>0, or Percent> variable.

The dialog box for the entry of cut-point values and labels is the same as for a Row\_Cut .

For example, if INCOME and TOTAL AID are continuous variables, then defining

```
INCOME as a By_Cut variable with cut-points 0.5, 29999.5, 49999.5,
```

and

```
TOTAL AID as an Average variable,
```

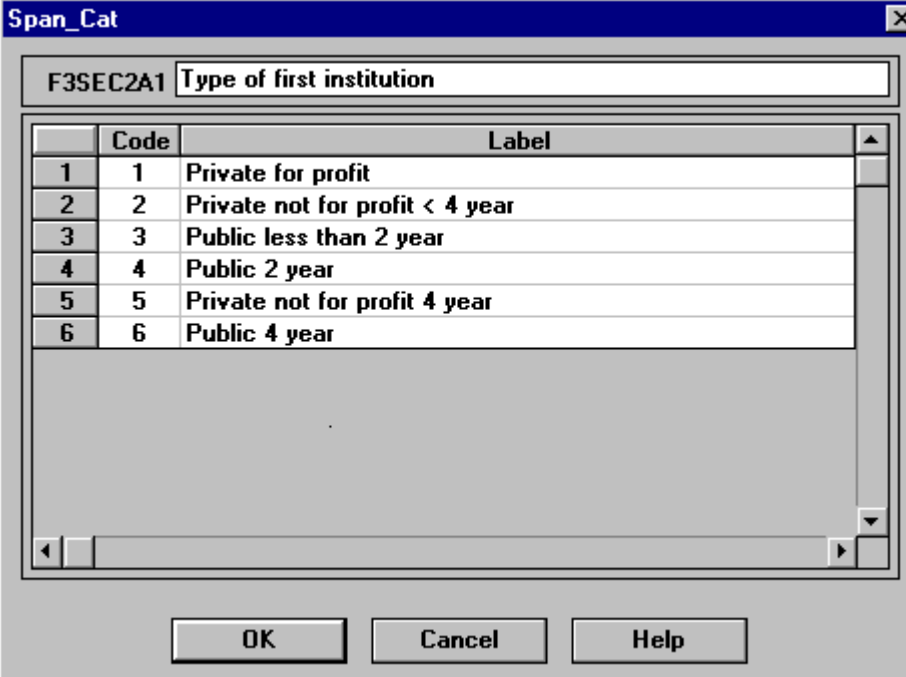
will produce a table with columns showing the average amount of aid received by aid recipients for incomes less than \$30,000, at least \$30,000 but less than \$50,000, and \$50,000 or more, as well as for all recipients (total).

See also: By\_Cat, By\_Lump, Row\_Cut.

# Spanner\_Cat

Spanner\_Cat uses **all** of the variable's categories to group sets of rows in the table. Each category will become a set of rows in the table. In the table presented in System Overview, the variable 'FIRST PSE Type' is a Spanner\_Cat.

The dialog box for the editing of Spanner\_Cat labels appears as follows:



The Span\_Cat dialog box is titled "Span\_Cat" and contains a text field for the variable name "F3SEC2A1" and a label "Type of first institution". Below this is a table with two columns: "Code" and "Label". The table lists six categories. At the bottom of the dialog are three buttons: "OK", "Cancel", and "Help".

	Code	Label
1	1	Private for profit
2	2	Private not for profit < 4 year
3	3	Public less than 2 year
4	4	Public 2 year
5	5	Private not for profit 4 year
6	6	Public 4 year

Users may edit the variable label and value labels. When you are satisfied with the labels, click on the OK button.

Warning: Spanner variables may drastically increase the number of estimates in a table. In some cases (with several rows and columns), the addition of a spanner variable may result in a table too large for the DAS to estimate.

# Spanner\_Lump

Spanner\_Lump produces subtables or sets of rows in a table. Lumping operates on categorical variables to define groupings of categories.

The dialog boxes for the entry of the lump specifications are the same as for a Row\_Lump.

See also: Spanner\_Cat, Spanner\_Cut , Row\_Lump.



# Spanner\_Cut

SPANNER\_CUT may be used to group sets of rows by creating categories from continuous variables or contiguous categorical variables. SPANNER\_CUTs are defined like ROW\_CUTs, only they form subtables or row groupings. The dialog boxes for the definition of Spanner\_Cut variables is the same as used for ROW\_CUTs.

See also: Spanner\_Cat, Spanner\_Lump, Row\_Cut.

# Filter

Filters subset the sample and operate using "select if" logic. Filters work with OR logic within a variable. When multiple Filter tags are used, the Filters operate with AND logic. You may use **ONE** Or\_Filter to get **or** logic between variables, but you may **NOT** do complex logic such as "(A AND B) OR (C AND D)".

You may use a variable more than once as a filter.

Filters may be integer filters (usually for categorical variables) or cut-point filters (usually for continuous variables).

## Integer Filters

You may enter an Integer Filter value(s) by clicking on the "Integer Filter" button at the top of the window. Each integer filter value must be a whole number, and cannot contain a decimal point. When an Integer Filter has several values, the Filter operates with OR logic to include all the values in the denominator of estimates.

The following dialog box appears for entry of Integer Filter labels and values:

Filter

☒ Integer Filter ☐ Cut-Point Filter

PCT4 Race/ ethnicity 1994

	Value	Label
1	1	Asian or Pacific Islander
2	2	Hispanic regardless of race
3	3	Black not of Hispanic origin
4		
5		
6		
7		
8		
9		
10		
11		

OK Cancel Help

For the table, only Asian, Hispanic, or Black individuals would be included in the table. The default labels are the variable label and value labels for the tagged variable. You may edit these labels as needed.

When you complete entering data, click on the OK button.

## Cut-Point Filter

You may enter a Cut-Point Filter by clicking on "Cut-Point Filter" button at the top of the window.

A cut-point value must be a fractional value (including a decimal point), not an integer. For example, 1.5, 2.5, and 99999.5 are all valid cut-point filter values, but 1.0, 1, and 10000 are not. Each cut-point filter must have ONE and only ONE value per tag.

For cut-point filters, a positive value means keep all values greater than the specified value. A negative cut-point filter means keep all values less than the specified value (including zero and -1 or missing). **WARNING:** *Be sure to exclude missing values (-1) if you do not want them when using a negative Cut-Point.*

The dialog box for the entry of Cut-Point Filter labels and values appears as follows:

Filter

☐ Integer Filter ☒ Cut-Point Filter

PCT4 Percent Asian, Fall 1992

**\*\* Floating point label allows only one filter. \*\***

Filter Value: 0.0

Filter Label:

OK Cancel Help

You may edit the variable label. You must enter the Cut-Point value in the Filter Value box (be sure to include a decimal point in the fraction). You must enter a label for the range of Filter values included in the Filter Label box.

### Example

To select Hispanic OR Black males with family income between \$20,000 and \$50,000, you would tag the following variables and specify as shown below:

```
Filter RACE2 Race/ethnicity
  3 Black
  4 Hispanic
Filter SEX Gender of student
  1 Male
Filter INCOME Adjusted gross family income
  19999.5 Income $20,000 or more
Filter INCOME Adjusted gross family income
  -50000.5 Income $50,000 or less
```

See also: Or\_Filter.

# Or Filter

Or Filters use "select if" logic in conjunction with Filters to specify subgroups. That is, to be included in a table an observation must either pass all of the Filters, **or** it must pass the Or Filter. This is true regardless of the order of the filters. You may use only one Or Filter with one or more Filters (which must be specified).

Or Filter works with OR logic in relation to the subgroup defined by the Filters. Or Filters may be integer filters or cut-point filters (see Filters for description of entry of Integer and Cut-point filters).

The DAS cannot perform complex filtering such as "(A and B) or (C and D)".

## **Example**

To select Hispanic or Black students OR all students with family incomes below \$20,000, you would tag the following variables and specify as shown below:

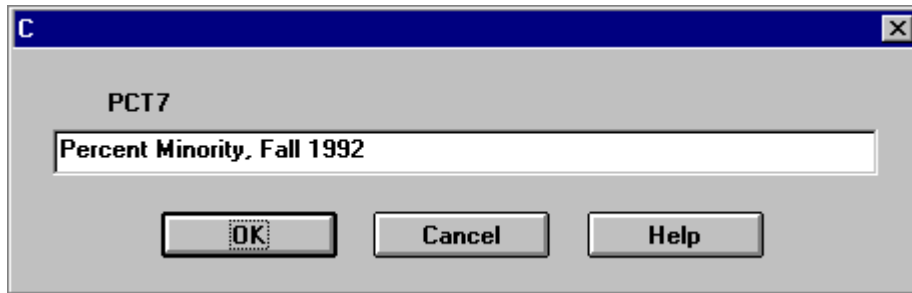
```
Filter INCOME Adjusted gross family income amount
-19999.5 Less than $20,000
OR_Filter RACE Race/ethnicity
3 Black
4 Hispanic
```

See also: Filter, Weight.

# Continuous

The Continuous menu option tag specifies a single continuous variable to be added to the correlation matrix. A continuous tag adds only one variable to the correlation matrix.

The dialog box for editing the label appears below:



Due to string length limits in SAS and SPSS, users should edit most DAS labels.

*WARNING: The only missing values specified for continuous variables is -1. Users must use Filters to further restrict the sample.*

# Cut

The Cut menu option tag creates dummy variable(s) {-1,0,1} for the correlation matrix from continuous variable ranges.

The dialog box for the entry of labels and values is as follows:

	Lower Cut	Upper Cut	Label
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

First, users should edit the variable label. The dummy variable label in the correlation matrix will have a label that is the concatenation of the variable label and the range label. Due to string limits in SPSS and SAS, users should edit the labels to shorten them.

Second, enter the maximum **Missing Value**. For example, a missing value of 24 will assign missing value status to all values in the -1, 0, 1, ... , 24 range. For many variables, the missing value will be the default -1, but for several others a value of 0 should be used.

Third, enter the **Lower Cut** value. This value must be a fractional number with a decimal point.

Fourth, enter the **Upper Cut** value. This value must be a fractional number with a decimal point.

Fifth, enter the **Label** for the defined range (i.e, values between the Lower Cut and Upper Cut values).

Repeat the third through fifth steps until all dummy variables are defined.

When finished, click the **OK** button.

## Example:

To create a dummy variable for marital status "divorced"

0 = Never married

1 = Separated

2 = Widowed  
3 = Divorced  
4 = Married

by define missing as 0, the lower cut point as 2.5 and the upper cut point as 3.5. A dummy variable will be created that has missing value -1 for values of -1 and 0; zero values for values of 1, 2, and 4; and one values for a value of 3. The appropriate label for this dummy is "Marital status=divorced".

*WARNING: Users may overspecify Cut dummy variables forming linear combinations within regression analyses. The DAS software does not check for overspecification.*

See also: Continuous.

# Each

The Each menu option creates a set of dummy  $\{-1,0,1\}$  variables from the values of the tagged variable. The variable must be categorical. The user is prompted for a base category, and dummy variables are created for each category of the variable as an offset from the base category. Each of the dummy variables are added to the correlation matrix.

When categorical variables have a reserved code of zero these values are treated as missing. To include the zero values, users must use a Lump.

Due to string length limits in SAS and SPSS, users should carefully edit all labels.

The dialog box for entry of values and labels is as follows:

The dialog box titled 'E' contains the following information:

Variable: PCT4 Race/ ethnicity 1994

Base Category Code: 1

	Code	Label
2	2	Hispanic regardless of race
3	3	Black not of Hispanic origin
4		
5		
6		
7		
8		
9		
10		
11		
12		

Buttons: OK, Cancel, Help

First, edit the variable label.

Second, select a **Base Category Code**. Users may click on the down arrow to get a list of valid values.

Third, edit the value labels.

When you are satisfied with the labels and selections, click the **OK** button.

## Example:

For PCT4 with values of

- 1 = White not of Hispanic origin
- 2 = Hispanic regardless of race



3 = Black not of Hispanic origin

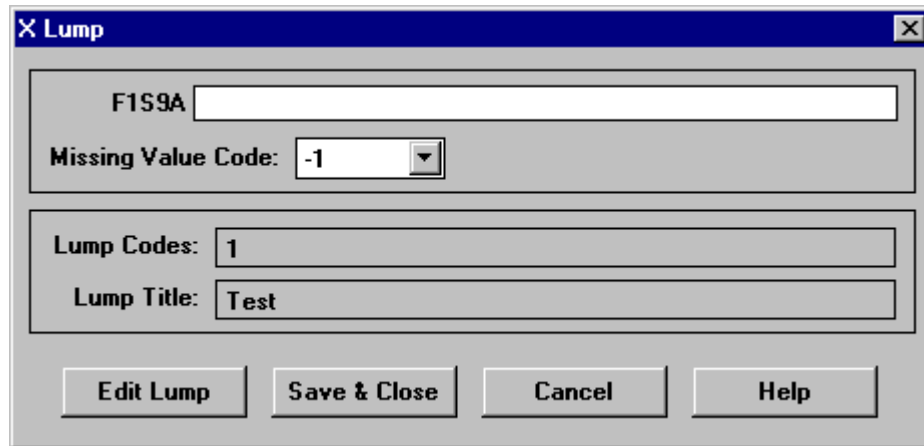
Selection of 1 as the Base Category Code will result in TWO dummy variables. The first should be labeled as "Race=Hispanic" and will have missing values when PCT4 has a value of -1, zero when PCT4 has a value of 1 or 3, and one when PCT4 has a value of 2. The second dummy should be labeled as "Race=Black" and will have missing values when PCT4 has a value of -1, zero when PCT4 has a value of 1 or 2, and one when PCT4 has a value of 3.

# Lump

A Lump is used to define one dummy (-1,0,1) variable for a correlation matrix. Groups of categories are selected to assign the value 1 vs. all other non-missing cases (assigned to zero). Users may define several Lumps for the same categorical variable.

For example, for race/ethnicity, grouping Hispanics, Blacks, and American Indians would produce a dummy variable with values of 1 for Hispanics, Blacks, and American Indians and values of zero for Whites and Asians.

The first of two dialog boxes to appear for a lump tag is as follows:



The screenshot shows a dialog box titled "Lump". It has a text input field containing "F1S9A". Below it is a "Missing Value Code" dropdown menu currently showing "-1". Further down are two more text input fields: "Lump Codes" containing "1" and "Lump Title" containing "Test". At the bottom of the dialog are four buttons: "Edit Lump", "Save & Close", "Cancel", and "Help".

First, enter a label for the lump variable.

Second, enter or edit the **Missing Value Code**. For many categorical variables with a reserved code of zero, the missing value should be changed to zero.

Third, click on the **Edit Lump** button. The second dialog box will appear as follows:

**X Lump**

F1S9A

Missing Value Code:

Lump Codes:

Lump Title:

**Enter Lump Row**

Categories:

- Code
- Once or twice**
- 5 - 10 days
- 11 - 15 days
- 16 - 20 days
- 21 or more

Lump Title:

Fourth, select the category or **categories** that should be assigned a value of one in the dummy variable.

Fifth, enter a label for the group of categories in the **Lump Title:** box.

Sixth, click the **Save Lump** button. The first dialog box will reappear.

Seventh, if you wish to edit the labels or the grouping click on the **Edit Lump** button. Otherwise, click on the **Save & Close** button.

*WARNING: With multiple Lumps, users may overspecify dummy variables, resulting in linear combinations within regression analyses. The DAS software does not check for overspecifications.*

See also: Each, Cut.

# Weight

Weight selects a variable to be used as the weight in the calculation of the estimates. All of the weights are labeled WEIGHT: ... and are shown in red. The DAS will not produce unweighted tables; if no weight variables exist, it means there is only one weight, which will be applied automatically.

If more than one weight is available, choose the weight which corresponds to the population you wish to describe in your table. If you are unsure of which weight to use, read the variable descriptions or check the NCES methodology reports at <http://nces.ed.gov>.

Weights can be used to restrict the data to a certain population subset or to ensure that the final output is representative of the population it is supposed to describe. For example, in NELS there are separate weights that describe the four populations represented:

- 1988 8th graders
- 1990 10th graders
- 1992 12th graders
- 1992 high school graduates.

Many cases within the DAS for NELS have zero weights--8th graders who dropped out and did not graduate in 1992 have zero weights for the 1992 high school graduates weight.

Another common use of weights is to account for nonresponse. For example, in the NPSAS:96 surveys, some questions (the CATI questions) were only asked of a small number of students. In order to be compared to other tables run from the same data, these respondents must have their weights adjusted upwards when CATI variables are used in the table. Therefore the NPSAS:96 data set has two weights: one for use with non-CATI variables and another for with CATI variables.

The dialog box for editing the label is as follows:

